

ABSTRACT

The present invention is a system, method, and computer program product for tracking and measuring digital content that is distributed on a computer network such as the Internet. The system collects online advertisement data, analyzes the data, and uses the data to calculate measurements of the prevalence of those advertisements. The system processes raw traffic data by cleansing and summarizing the traffic data prior to storing the processed data in a database.

An advertisement sampling system uses site selection and definition criteria and a probe map to retrieve Web pages from the Internet, extract advertisements from those Web pages, classify each advertisement, and store the data in a database. A statistical summarization system accesses the processed raw traffic data and the advertisement data in the database to calculate advertising prevalence statistics including the advertising frequency, impressions, and spending.